Yolo County Probation Department

Ohio Risk Assessment System Pretrial Assessment Tool Validation Study

Yolo County Probation

Ohio Risk Assessment System Pretrial Assessment Tool (ORAS-PAT) Validity Study

This report was developed by RDA Consulting under contract with the Yolo County Probation Department.

RDA Consulting, April 2024

Contents

Introduction

Applying the principles of effective intervention, or the risk, need, responsivity (RNR) model in corrections, reduces the likelihood that an individual will reoffend.[1] The RNR model is best executed through validated risk/needs assessment (RNA) tools. Correctional risk and needs assessments have advanced in recent years and are currently in their fourth generation, which allows for the measurement of static and dynamic factors that not only predict risk but identify need areas and can be used to help determine supervision and service needs. In 2010, the University of Cincinnati Corrections Institute (UCCI) developed the Ohio Risk Assessment System (ORAS) in collaboration with the Ohio Department of Rehabilitation and Correction as a statewide system to measure individuals' risk and needs while considering the points in the system in which the assessment is taking place. To this end, several tools spanning the criminal justice system were validated, including the Pretrial Assessment Tool (PAT).[2]

Since initial testing, the PAT has been adopted in more than 50 jurisdictions nationwide to predict a defendant's likelihood of failing to appear in court (FTA) or committing a new offense while released pretrial. However, when applied to jurisdictions outside Ohio (the development sample), instruments become vulnerable to performance or prediction shrinkage (declines in predictive accuracy or validity) as individuals' characteristics, laws, agency policies, and local supervision conditions inevitably vary from jurisdiction to jurisdiction.

Ultimately, RNAs cannot be valid if they are not reliable at the local level. Additionally, RNAs must be fair to all assessed subpopulations. This means instruments should evaluate racial and ethnic minorities and socioeconomically disadvantaged individuals equitably to ensure all groups receive supervision and treatment that corresponds with their actual levels of risk.

In compliance with Senate Bill 36 (SB 36), Yolo County Probation Department's use of the PAT was locally validated in 2021 and found to predict pretrial failures accurately. This previous validation study did not observe any statistically significant differences in risk levels across gender, race, ethnicity, or offense type. To adhere to risk assessment validation best practices and SB 36's requirements, it is important to re-evaluate the PAT's local predictive validity regularly.

The ORAS PAT in Yolo County

Recognizing advancements in risk assessment research and the evolution of fourth-generation tools, Yolo County adopted the ORAS in 2010 to help its Adult Services and Court staff make more informed decisions at critical points in the criminal justice system process.

Probation Officers in the Department's Pretrial Services Unit administer the PAT for every eligible defendant held in custody to determine the suitability of release before sentencing and general supervision during their pretrial release.[3] The PAT contains seven items: (1) Age at first arrest; (2)

---

[1] Bonta, J., & Andrews, D. (2007). Risk-need-responsivity model for offender assessment and rehabilitation. Ottawa, ON: Public Safety Canada.

[2] Latessa, E. J., Lemke, R., Makarios, M., Smith, P., & Lowenkamp, C. T. (2010). The creation and validation of the Ohio Risk Assessment System (ORAS). Federal Probation, 74(1), 16–22.

[3] According to the Yolo County Pretrial Services Manual, defendants held in custody are not eligible for pretrial release if they meet any of the following conditions: (1) has to register pursuant to Penal Code Section 290, (2) currently supervised on Post Release Community Supervision (PRCS), (3) currently supervised on Mandatory Supervision, (4) currently has an out-of-county felony warrant, (5) in custody for a fresh offense that is violent, pursuant to Penal Code Section (§) 667.5, (6) in custody for a fresh offense that is serious, pursuant to Penal Code § 1192.7, (7) has a "No Bail" in the amount of bail.

Number of failure to appear (FTA) warrants in the past 24 months; (3) Number of prior jail incarcerations; (4) Employment status at the time of arrest; (5) Residential stability; (6) Illegal drug use during the past six months; and (7) Severity of drug use problem. The PAT is scored by summing item responses to achieve a total score corresponding with a "low," "moderate," or "high" risk score. Cutoff scores for the tool are displayed in **Table i** below. The PAT is scored with information collected in a short face-to-face interview and verified with collateral information from other records (e.g., Probation notes and FTA warrants) and sources (e.g., family, friends, victim).[4] UCCI provides a scoring guide for the PAT that factors in the responses to the interview questions.

Table i. PAT Total Score Cutoffs

Risk Category	Total Score
Low	0-2
Moderate	3-5
High	6-9

Officers provide the Court with PAT assessment results in a written report, including scoring for individual items. Additional information regarding the defendant is also provided, such as booking charges, criminal conviction history, substance use and mental health history, whether contact has been made with a victim, and, when applicable, Ontario Domestic Assault Risk Assessment (ODARA) results.[5] Pretrial Officers consider these additional pieces of information, along with PAT results, as part of a release decision framework when making a release recommendation to the Court.

Ultimate release decision authority rests with the Judicial Officer of the Court, who can deny pretrial release, release defendants on their own recognizance, or release defendants under the supervision of Probation's Pretrial Services Unit. Individuals released on their own recognizance are required to appear in Court but typically do not have any Probation reporting requirements or other terms for their release. When a defendant is not deemed appropriate to be released on their own recognizance by the Judicial Officer, defendants are placed on supervised release with terms and conditions (e.g., office visits and GPS monitoring) to ensure they appear in Court.

Current Study

The Yolo County Probation Department (the County) contracted with RDA in October 2023 to complete a validation of its pretrial risk assessment tool that will be shared with the public to comply with the requirements of SB 36. The goals of the project are to determine if the tool is accurately predicting defendants' risk to recidivate (i.e., failure to appear in court, charges or arrests for new offenses, and pre-trial release revocations) and understand and reduce any detectable biases based on gender, race, or ethnicity in pretrial release decision-making. Specifically, the following objectives guided the study design:

Objective 1. Learn about local assessment policies and practices through survey data collection with probation officers.

---

[4] (2015). Ohio Risk Assessment System: Pre-Trial Assessment Tool (PAT). *University of Cincinnati Corrections Institute*, 7, 3-4. https://www.yolocounty.org/home/showpublisheddocument/69979/637605641001200000. Pretrial Services Unit Manual. *Yolo County Probation Department*.
[5] In cases of domestic violence, officers incorporate Ontario Domestic Assault Risk Assessment (ORDARA) tool results in their release recommendation.

Objective 2. Investigate predictive validity or the extent to which the PAT correctly measures what it is supposed to measure—in this instance, the likelihood of a pretrial failure.

Subsequent sections of this report define the methods used to meet these objectives and describe the study findings. Section II presents local assessment practices and reports predictive validity findings before a final summative discussion of the study's findings and recommendations for instrument and agency improvement.

I. Methodology

To evaluate the effectiveness of the PAT, RDA followed the data collection, sampling, and analysis methodologies outlined in this section. This study employed a mixed-methods approach to data collection and analysis, integrating staff perspectives with quantitative rigor to paint a complete picture regarding assessment practices and implementation while contextualizing predictive validity results to provide more precise recommendations.

Local Assessment Practices

To understand local assessment practices and how the PAT is used in the Department, RDA developed and administered officer surveys to explore the implementation of the tools, including whether assessors are conducting interviews as part of the risk assessment and how the results are used to inform release recommendations and decisions. Additionally, RDA spoke with Probation administrators to understand department policies and day-to-day practices. Qualitative data findings from the survey and interview and quantitative findings from officer surveys help RDA and the County identify facilitators and barriers to PAT assessment implementation and administration.

Recruitment. Yolo County Probation administrators identified nine officers in the Pre-Trial Services Unit who had experience administering the PAT. Probation administrators sent an introductory e-mail to the selected survey participants informing them about the evaluation and when to anticipate an e-mail from RDA with a unique survey link. RDA initiated outreach on November 15, 2023, and collected responses from each of the nine officers by November 30, 2023.

Questions. RDA developed a survey tool with input from Probation administrators to understand how the PAT is conducted and identify the tool's strengths and challenges. The complete survey tool can be reviewed in **Appendix A**. Specific questions included how PAT results are being properly applied and whether assessors are conducting interviews and referring to scoring guides when utilizing the instrument. RDA used Alchmer, a secure, web-based survey platform, to administer and collect survey responses from officers.

Analysis. Qualitative data collected from open-ended officer responses were analyzed using directed content analysis strategies to identify and code themes identified by common words and similar explicit or inferred meanings within participants' responses. Additionally, responses on themes were grouped to allow for some basic quantification (e.g., this many respondents referred to a particular "strength" or "limitation" when administering the PAT or calculating assessment results). Quantitative data produced from the survey were analyzed to produce descriptive statistics such as frequencies and averages.

PAT Predictive Validity

To evaluate the predictive accuracy of the PAT, RDA collected administrative data from Yolo County related to defendants who received assessments, their assessment scores, and outcomes. The predictive validity sample spanned four years and included defendants with PAT assessments completed between 2020 and 2023. Analyses related to the evaluation of the psychometric properties of the instruments and normative comparisons helped RDA determine whether PAT items are theoretically relevant and connected with the measurement of criminogenic risk. Units of analysis were defendants held in custody and released before sentencing.

Sampling Strategy. The Yolo County Probation Department provided data with defendants' names redacted for this validation study. Specifically, Probation provided three data files,

including PAT results, actions data (e.g., release decisions and pretrial failures), and Probation release recommendations. The evaluation team joined these datasets with "Person ID" fields that identified individuals across all files provided, matching PAT results for 1,863 distinct assessed defendants to the nearest pretrial release action (i.e., denial or pretrial release decision within 180 days of an assessment).[6] Ultimately, 103 distinct assessed defendants did not have a near-term pretrial release action, reducing the validation sample size to 1,760 distinct defendants.

After matching assessment and action data, the evaluation team selected each defendant's earliest assessment for analyses to align with statistical modeling assumptions and provide the longest follow-up period (i.e., time at risk for recidivism).[7] The evaluation team then excluded defendants without the requisite fixed three-month follow-up duration from the sample (i.e., not denied or released pretrial before October 1, 2023), further reducing the validation sample size to 1,622 distinct defendants.

Lastly, the validation sample was limited to defendants released pretrial, including defendants on supervised release and defendants released on their own recognizance.[8] About 45% of all assessed defendants were ultimately released pretrial before October 1, 2023, including 45% of male and 47% of female defendants. Additionally, 47% of White and Hispanic defendants were released, 41% of defendants of an "Other" race, and 39% of Black defendants. **Differences in pretrial release rates did not rise to the level of statistical significance for gender or race, meaning that the observed differences may be due to chance.[9]** See **Appendix B** for a detailed breakdown of defendant characteristics by release decision.

Ultimately, the department completed PAT assessments for 729 distinct released defendants between 2020 and 2023 based on the exclusion criteria described in the preceding paragraphs. Displayed in **Table 1** are the characteristics of these 729 released PAT defendants included in the predictive validity study sample.[10] Sample sizes must be large enough to support cross-tabulations of discrete assessment groups and subgroups of interest within each assessment sample. If too few cases are collected, the number of cases in analytical subgroups will fall below levels supporting meaningful analysis. In short, larger samples will support more reliable and informative findings.[11]

Predictor & Outcome Variables. Predictor variables for this study include the results of a defendant's earliest matched assessment, including, for example, the total risk score, item scores, and overall risk level. Dichotomous recidivism outcome measures, including (1) failure to appear in court (FTA) as required, (2) arrest for a new offense during the pretrial period, (3) pretrial release

---

[6] Three-quarters of assessments identified as a "nearest match" were completed within four days of a release decision.

[7] For the purposes of validation analysis by charge level, the research team also retained each defendant's most serious charge (i.e., felony and enhancements).

[8] Regardless of their supervision status, all defendants released pretrial could have "failed" by committing a new offense or not appearing in court. This study's goal is to assess the PAT's ability to correctly predict any pretrial failure, not the effectiveness of supervision in preventing a failure. Therefore, all released defendants meeting the study exclusion criteria for follow-up duration are included in the analysis sample.

[9] Approximately 47% of White and Hispanic defendants were released followed by 41% of defendants of an "Other" race, and 39% of Black defendants. Approximately 45% of male defendants were released and 47% of female defendants.

[10] Gender was missing for 5 released defendants, equivalent to .7% of the total 729 defendants with a PAT assessment based on the exclusion criteria used in this study. Additionally, charge level was missing for 25 defendants, equivalent to 3.4% of the total defendants with a PAT assessment based on the exclusion criteria. Individuals with missing gender or charge level were not examined for sub-analyses by gender identity or charge level due to the small samples available.

[11] Mamalian, C. A. (2011). State of the science of pretrial risk assessment. Jointly published by the Pretrial Justice Institute and the Bureau of Justice Assistance. Washington D.C. Retrieved from https://bja.ojp.gov/sites/g/files/xyckuh186/files/Publications/PJI_PretrialRiskAssessment.pdf

revocation, and (4) any failure (i.e., FTA, new offense, or revocation), were created to be used in determining the predictive accuracy and fairness the PAT.[12] The evaluation team examined outcomes within a three-month follow-up period of release, the sample average length of pretrial release. A shorter follow-up period based on a defendant's time in the community during the pretrial stage applies to the PAT.

Table 1. Defendant Sample Characteristics (*N* = 729)

Characteristic	*n*	% or $\bar{x}$ (SD)
Gender		
Male	583	80.5%
Female	141	19.5
Race		
White	301	41.3%
Hispanic	251	34.4
Black	120	16.5
Other	57	7.8
Age Cohorts		
18-26	138	18.9%
27-34	209	28.7
35-44	228	31.3
45-54	89	12.2
55-74	65	8.9
Age	729	36.6 (11.2)
Most Serious Booking Charge Level		
Misdemeanor or Infraction	68	9.7%
Felony or Enhancement	636	90.3
Risk Level		
Low	136	18.7%
Moderate	273	37.4
High	320	43.9
PAT Total Score	729	4.8 (2.3)

Statistical Analyses. All data underwent a preliminary round of data preparation in Excel before importing into Stata for statistical analyses. First, the team calculated univariate descriptive statistics for defendant demographic characteristics presented in **Table 1** along with justice processing measures to describe the nature and distribution of the PAT assessment sample.

Second, the evaluation team normed the PAT instrument based on assessments completed in Yolo County. The norming process involved establishing the minimum and maximum total risk and item scores and calculating frequencies to examine the distribution of defendants falling into each risk level category. The team also calculated average total and item scores for the

---

[12] FTA included the following actions: Pretrial Terminated/ FTA to Court, Warrant Issued, Warrant Served, Warrant Served/ Recalled; Arrest for a new offense included the following action: Pretrial Terminated/ New Offense; Revocation failure included the following actions: Pretrial Terminated/ Non-Compliance, Pretrial Terminated/ GPS Non-Compliance, Pretrial Terminated/ SCRAM Non-Compliance.

complete sample of assessments and executed t-tests to examine possible differences in risk and need as a function of each subgroup of interest (e.g., race or ethnicity, gender).

To explore the internal reliability of the PAT, the evaluation team calculated Cronbach's alpha internal consistency coefficients. This measure determines the level of internal consistency between item scores to determine how closely related they are as a group. A coefficient score greater than or equal to .70 is typically considered acceptable. Additionally, AUC-ROC (area under the receiver operating characteristic curve) was calculated and interpreted using the following minimum values: fair (AUC-ROC = .55), good (AUC-ROC = .64), and excellent (AUC-ROC = .71).[13] Of note, the AUC-ROC is not affected by samples' base rates of reoffending and is therefore considered a better measure of internal reliability than Pearson's *r* correlation coefficient.

The evaluation team examined the predictive validity of the PAT with basic frequencies and chi-square tests of significance to examine the relationship between assessed risk levels and outcomes (i.e., FTA, new offense, revocation, and "any pretrial failure"). The team also utilized Pearson's *r* to examine the direction and relative strength of the relationship between assessment scores and recidivism. Pearson's *r* was interpreted using the following minimum values: fair (*r* = .10), good (*r* = .24), and excellent (*r* = .37).[14] Finally, the evaluation team used AUC-ROC and logistic regression to measure the PAT's ability to predict pretrial failure correctly. While investigating predictive validity for subgroups of interest, the research team limited pretrial outcomes to "any pretrial failure" to produce more informative findings with the largest sample size possible.

---

[13] Desmarais, S. L, Singh, J. P. (2013). Risk assessment instruments validated and implemented in correctional settings in the United States. Lexington, Kentucky. *Council of State Governments, 2.* Accessed here: *https://csgjusticecenter.org/wp-content/uploads/2020/02/Risk-Assessment-Instruments-Validated-and-Implemented-in-Correctional-Settings-in-the-United-States.pdf*
[14] Desmarais, S. L, Singh, J. P. (2013). Risk assessment instruments.

II. Results

Compiling data with a mixed-methods approach, **the RDA study team found that the PAT has excellent predictive validity for any pretrial failure in Yolo County**. The tool's predictive validity is promoted by local assessment practices that support accurate scoring and implementation fidelity. This section shares study results related to local assessment practices and predictive validity findings, informing recommendations that follow.

Local Assessment Practices

RDA collected survey responses from nine Probation Department Officers in the Pretrial Services Unit who are involved with administering the PAT to learn about local tool implementation. Additionally, RDA spoke with Probation administrators to understand department policies and day-to-day practices. As detailed in the preceding **Methodology** section, the evaluation team coded survey responses to identify facilitators and barriers to PAT assessment implementation and administration. Direct quotations in this report were edited for clarity. The research team also incorporated descriptive statistics from the predictive validity data sample to further contextualize survey findings.

The Pretrial Services Unit demonstrated several strengths in their local administration of the PAT. All officers have been trained to use and administer the tool according to best practices: meeting face-to-face with defendants, referring to the interview guide, and collecting collateral information to inform their release recommendation to the Court according to their release decision framework. While officers do not always reference the scoring guide for the seven-item tool, supervisors review officers' PAT scoring and results. Additionally, officers are supportive of the PAT. Officers shared that the tool is easy to use and understand, produces consistent results, and generates valuable information that informs release recommendations while accurately predicting pretrial release failures.

This report shares relevant qualitative and quantitative findings exploring local PAT assessment practices in greater detail for the following areas: (1) Administration fidelity; (2) Release Recommendations; (3) Officer buy-in.

Administration Fidelity

Although the PAT is simple, with just seven items, all officers surveyed reported they were formally trained to use the PAT, and almost half had participated in a booster or recertification training. In keeping with UCCI and department guidance, all officers also reported conducting interviews with the defendant during the PAT assessment. Additionally, most officers (78%, $n = 7$) use the PAT Interview Guide during interviews. All officers also consult collateral information, most frequently referring to justice records such as RAP sheets, prior Probation notes, and FTA warrants. Four officers (44%) stated they also consult defendants' family and friends.

Just three officers (33%) stated they reference the scoring guide when scoring a PAT assessment. The largest share of officers (56%, $n = 5$) reference the scoring guide on an "as needed" basis, e.g., if they have a question about a particular answer. Ultimately to ensure accurate scoring, almost all officers (89%, $n = 8$) reported that there is a quality assurance process in place for reviewing PAT results (e.g., supervisor approval).

Overall, officers shared limited challenges related to PAT administration. When officers did report implementation barriers, they were most often related to defendants declining to be interviewed

or providing responses contradicted by collateral information. Officers also shared difficulties when conducting assessments with defendants who have severe drug abuse problems or alcohol addiction. According to officers, these defendants may be falling asleep, lacking focus, or otherwise in active detox during their interviews.[15]

Release Recommendations

Following the local release decision framework, all officers reported that PAT results and other information collected during their pretrial release investigation inform release recommendations to the Court. When making recommendations, officers most often reference the following pieces of collateral and additional information to corroborate defendant responses and complete release investigations:

- Victim statements and concerns;
- Previous justice records and contacts (e.g., criminal record history, previous FTA warrants);
- Seriousness of the current (i.e., instant) offense; and
- Domestic violence[16]

While all factors may be considered, officers shared that victim statements and the severity of the current offense are prioritized when making a release recommendation to the Court. For example, according to state law (Cal. Pen. Code § 1270.1), Probation cannot recommend pretrial release for certain violent or serious offenses until a public hearing has been held.

As mentioned previously in the **Introduction**, the ultimate release decision authority in Yolo County rests with the Judicial Officer of the Court, who can deny pretrial release, release defendants on their own recognizance, or release defendants under the supervision of Probation's Pretrial Services Unit. To better understand how Probation's recommendations may differ from the decision of the Judicial Officer, the RDA study team incorporated Probation release recommendations into the analysis sample prepared for the validation study. Of note, release recommendation data was only available between July 2022 and December 2023.

While Probation was largely in agreement with the Judicial Officer's choice to deny pretrial release (agreeing in 91% of cases), the Judicial Officer disagreed with Probation's recommendation in 65% of cases when releasing defendants (see **Table 2**).[17]

Table 2. Probation Release Recommendation Agreement with Judicial Officer, by Pretrial Release Decision

	Overall (*n* = 319)		Denied (*n* = 158)		Released (*n* = 161)	
	n	%	*n*	%	*n*	%
Agreement Type***						
Agree	201	63.0%	144	91.1%	57	35.4%
Disagree	118	37.0	14	8.9	104	64.6

* p < .05, ** p < .01, *** p < .001

---

[15] Very rarely, defendants with severe drug abuse problems may be in the jail medical facility and therefore not available for interview. Interviews are not required as part of the PAT assessment; however, officers make at least one other effort to meet with defendants who might have been in the jail medical facility prior to submitting any report.
[16] In cases of domestic violence, officers incorporate Ontario Domestic Assault Risk Assessment (ORDARA) tool results in their release recommendation.
[17] Analysis is limited to those defendants with available Probation release recommendation data. Refer to **Appendix C** for detailed defendant characteristics by Probation release recommendation agreement with the Judicial Officer.

Probation may have recommended denying pretrial release in cases where a defendant is ultimately released if a victim statement was not provided at the time of their report to the Court but was eventually contacted before the defendant's hearing. Probation's recommendation to deny pretrial release may also conflict with the Court's release decision if the defendant was charged with certain violent or serious offenses. As mentioned previously, Probation is required to recommend denying pretrial release in these cases, even if defendants may otherwise be a candidate for supervised release based on the results of the PAT, for example. Probation will inform the Judicial Officer during pre-arraignment that they initially recommended denying release because of the related state law (Cal. Pen. Code § 1270.1) but will accept the defendant on supervised release with the completion of the public hearing.

Officer Buy-In

Officers reported high levels of buy-in with the PAT. Almost all officers (89%, *n* = 8) rated themselves as satisfied with the PAT, including seven officers who shared that they are "very satisfied" with the tool.[18] Importantly, all officers are confident that a defendant would receive the same PAT score regardless of the officer member conducting the assessment.[19] Furthermore, all officers agreed that the PAT helps to determine who is appropriate for pretrial release.[20]

On average, officers <u>strongly agreed</u> with the following statements related to the administration of the PAT: [21]

- "The ORAS-PAT has helped our agency establish a common language regarding defendants' risk level."
- "The ORAS-PAT is easy to read, interpret, and use."
- "The ORAS-PAT increases consistency in rules, policies, and procedures regarding assessment practices."
- "There are clear guidelines for when to use the ORAS-PAT."
- "There is a protocol for how to use the ORAS-PAT."

On average, officers <u>agreed</u> with the following statements: [22]

- "The ORAS-PAT reduces variation in decision-making across individual officers."
- "The ORAS-PAT helps our agency make proper decisions regarding whether a defendant is appropriate for pretrial supervision versus custody."
- "The ORAS-PAT reduces bias in criminal justice decision-making in our agency."
- "Risk and needs assessment benefits defendants."

---

[18] Only one staff member gave a neutral rating for their satisfaction with the ORAS-PAT.

[19] Two officers (22%) are <u>completely</u> confident that a defendant would receive the same ORAS-PAT score regardless of the staff member conducting the assessment.

[20] Three officers (33%) <u>strongly</u> agree the ORAS-PAT helps determine who is appropriate for pretrial supervision.

[21] Officers were asked to share their level of agreement with the statements on a scale of "strongly disagree" (i.e., a score of one) to "strongly agree" (i.e., a score of five). The average score for "The ORAS-PAT has helped …" was 4.8, which corresponds with a ranking of "strongly agree." The average score for "The ORAS-PAT is easy …" was 4.7, which corresponds with a ranking of "strongly agree." The average score for "The ORAS-PAT increases consistency …" was 4.6, which corresponds with a ranking of "strongly agree." The average score for "There are clear guidelines …" was 4.7, which corresponds with a ranking of "strongly agree." The average score for "There is a protocol …" was 4.7, which corresponds with a ranking of "strongly agree."

[22] The average score for "The ORAS-PAT reduces variation …" was 4.1, which corresponds with a ranking of "agree." The average score for "The ORAS-PAT helps our agency make proper decisions …" was 4.3, which corresponds with a ranking of "agree." The average score for "The ORAS-PAT reduces bias …" was 4.1, which corresponds with a ranking of "agree." The average score for "Risk and needs assessment …" was 4.1, which corresponds with a ranking of "agree." The average score for "The ORAS-PAT assessment has made my job easier" was 4.1, which corresponds with a ranking of "agree." The average score for "There is staff support …" was 4.2, which corresponds with a ranking of "agree."

- "The ORAS-PAT assessment has made my job easier."
- "There is officer support for implementing the ORAS-PAT."

Overall, these findings indicate that the Pretrial Services Unit views the PAT as an easy tool to use and understand, producing consistent results that help generate release recommendations and accurately predict pretrial release failures. A representative quote is shared below:

"I believe the ORAS-PAT assessment is a valuable tool in evaluating the defendant for possible pretrial release. While the tool is not perfect, it provides another measure to evaluate the defendant." – Pretrial Services Unit

Predictive Validity

To examine the internal reliability and predictive accuracy of the PAT in Yolo County, the RDA evaluation team normed the instrument on the local population and executed several predictive validity analyses to measure the extent to which the tool accurately predicts recidivism. These findings establish the distribution of assessment scores for the evaluation period and whether PAT items are theoretically relevant and connected with the measurement of criminogenic risk of local defendants. As described in greater detail in the preceding **Methodology** section, these analyses were performed using administrative data from Yolo County related to defendants who have received assessments, their assessment scores, and outcomes. Additionally, the analysis was limited to a period of three months following pretrial release, the sample average length of release.

Most released defendants in the predictive validity sample (69%, n = 500) did not have an FTA, new offense, or revocation within three months of their pretrial release date (i.e., were successful, see **Appendix D** for a detailed distribution of pretrial successes and failures within three months for different sub-group of interest). About 91% (n = 124) of defendants assessed as low-risk were successful, while 77% (n = 209) of moderate-risk were successful, and 52% (n = 167) of high-risk were successful. These substantial differences across risk levels conform with our expectations for the PAT's performance. Other analyses presented in this section confirm that the PAT has excellent predictive validity for any pretrial failure in Yolo County.

This section reviews complete internal reliability and predictive validity findings in the following structure: 1) Norms & Internal Reliability; 2) Full Sample Predictive Validity; and 3) Race, Gender, and Charge Level Predictive Validity.

Norms & Internal Reliability

Defendants across the full PAT sample scored slightly less than a five (4.8, moderate risk) on average, as displayed in **Table 4** below. Male defendants (5.9, moderate risk) scored slightly higher than females on average (4.6, moderate risk); however, this average score difference does not rise to the level of statistical significance.[23]

White defendants (average total score 5.3, moderate to high risk) scored higher than all other non-White defendants on average (Hispanic average 4.5, moderate risk; Black average 4.6, moderate risk; defendants of an "Other" race average 4.2, moderate risk). This difference in average total scores between defendants of different races and ethnicities is statistically

---

[23] At the individual item level, average differences between males and females rose to the level of statistical significance for item 3 (p < .001).

significant. [24] Additional post-hoc analysis with both Tukey HSD and Tukey-Kramer methods revealed that White defendants in particular had statistically significant higher average total scores, item six scores (recent illegal drug use), and item seven scores (severity of drug use problem) than Black defendants and defendants of an "Other" race or ethnicity.[25] This difference in total average scores may be attributable in part to the older age of White defendants in the sample.[26]

Table 4. PAT Average Total Score & Internal Reliability Coefficients for Overall Sample & Subgroups, 3 months

	Overall (N = 729)	Male (n = 583)	Female (n = 141)	White (n = 301)	Hispanic (n = 251)	Black (n = 120)	Other (n = 57)
Average Total Score	4.8	4.9	4.6	5.3***	4.5***	4.6***	3.9***
Alpha	.63	.64	.59	.57	.66	.59	.65

*p < .05, ** p < .01, *** p < .001*

The internal reliability alpha coefficients for the PAT were below the .70 internal reliability threshold for all items in the overall sample and sub-analyses by gender and race (see **Table 4**). The alpha coefficients are used to calculate how closely a set of RNA items are related to each other. As a note, the alpha coefficient is sensitive to both the number of items and the distribution of item scores for a given tool. Because the PAT only has seven items scored between zero and two, the variability is limited, making the tool highly sensitive to any otherwise minor differences in scores.

To further understand the distribution of PAT scores, frequencies by risk level for the overall sample, gender, and race are displayed in **Table 5**. Findings show that slightly less than half of defendants were assessed as high-risk (44%), while approximately 37% were deemed moderate-risk, and the remaining 19% were deemed low-risk. In addition to having similar average PAT scores, male and female defendants were assessed as low, moderate, and high-risk in roughly similar proportions. A higher percentage of White defendants fell into the high-risk category (53%) in comparison to Hispanic (38%), Black (40%), and "Other" (30%) defendants. Meanwhile, defendants of an "Other" race fell into moderate (44%) and low-risk categories (26%) at the highest rate.

Table 5. PAT Distribution for Overall Sample & Subgroups by Risk Level

Risk Level	Overall (N = 729) n	Overall (N = 729) %	Male (n = 583) n	Male (n = 583) %	Female (n = 141) n	Female (n = 141) %	White (n = 301) n	White (n = 301) %	Hispanic (n = 251) n	Hispanic (n = 251) %	Black (n = 120) n	Black (n = 120) %	Other (n = 57) n	Other (n = 57) %
Low	136	18.7%	108	18.5%	27	19.1%	37	12.3%	61	24.3%	23	19.2%	15	26.3%
Moderate	273	37.4	212	36.4%	59	41.8%	104	34.6%	95	37.8%	49	40.8%	25	43.9%
High	320	43.9	263	45.1%	55	39.0%	160	53.2%	95	37.8%	48	40.0%	17	29.8%

Full Sample Predictive Validity

The following figure and table display the rates of any pretrial failure (i.e., FTA, new offense, or revocation) at three-month follow-up by PAT risk level for the full sample of defendants (see Figure

---

[24] At the individual item level, average differences between defendants of different race/ethnicities rose to the level of statistical significance for every item except item 4 (item 1, $p < .05$; item 2, $p < .01$; item 3, $p < .01$; item 6, $p < .001$; item 7, $p < .001$).

[25] Post-hoc analysis with both Tukey HSD and Tukey-Kramer methods revealed that White defendants scored higher than Black defendants and defendants of an "Other" race/ethnicity, with statistical significance ($p < .05$) on the following: total score, item 6, item 7.

[26] The average age of defendants in each racial group is as follows: White (38.1 years), Hispanic (34.5 years), Black (34.0 years), and "Other" race (35.0 years). These differences in age by racial group are different at a statistically significant level ($p < .001$).

1 and Table 6). Defendants assessed in the low-risk category had a 9% rate of any pretrial failure within the three-month follow-up period. Defendants in the moderate-risk category had a 23% pretrial failure rate within the same time frame, while defendants assessed in the high-risk category had a 48% pretrial failure rate. As shown in **Figure 1**, this step-patterned increase in pretrial failure rates as the risk level increases from low to high aligns with performance expectations.

Figure 1. PAT Any Failure Rate, by 3-Month Follow-Up Period & Risk Level *(N = 729)*

Any Pretrial Failure
$\chi2 = 80.2^{***}$
AUC = .72

* p < .05, ** p < .01, *** p < .001
n values reflect the number of defendants who failed in each category.
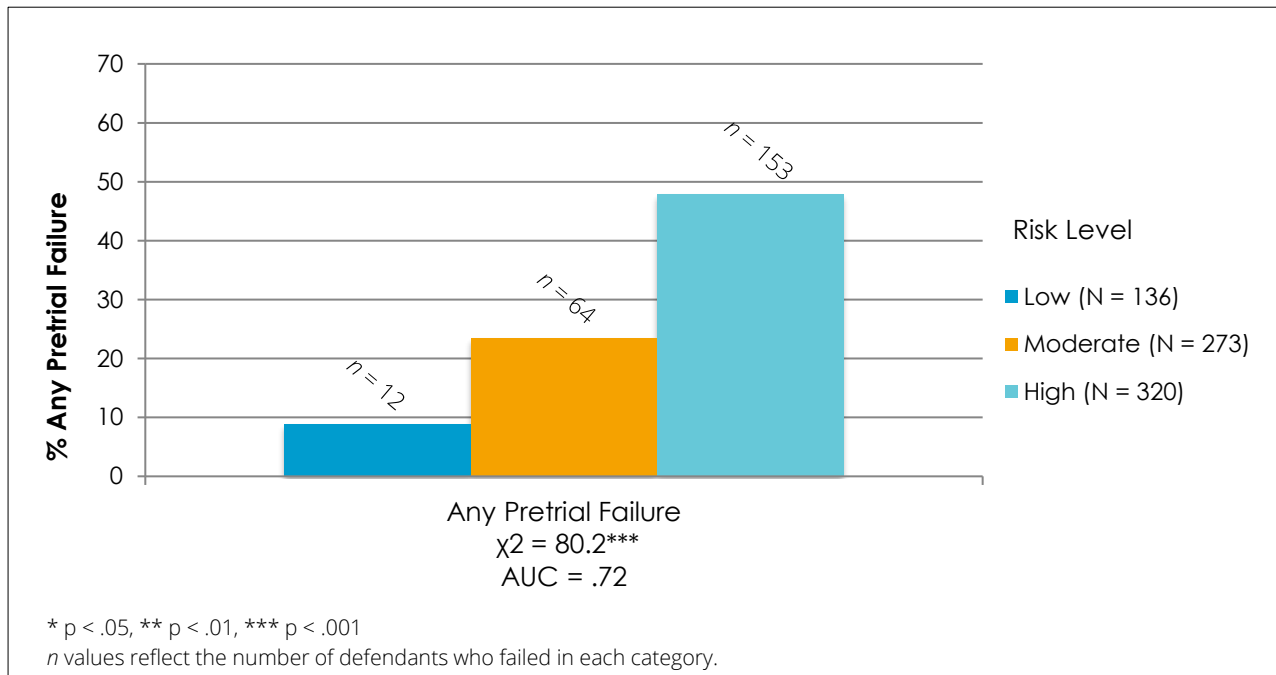
Additional testing displayed in **Table 6** found a statistically significant difference between risk levels and rates of any pretrial failure.[27] Confirming **Figure 1**'s visual trend of increasing pretrial failure rates moving from low- to high-risk, the calculated Pearson's *r* correlation coefficient showed a good-excellent predictive value (.34) that was statistically significant and positive, meaning that the chance of any pretrial failure increases as a defendant's assessed risk level increases.[28] This validation analysis found an excellent AUC-ROC value (.72) for the PAT overall sample. This finding indicates that a defendant referred again to the Pretrial Services Unit three months after release would have a higher PAT score compared to a randomly selected defendant not referred again about 72% of the time.

Table 6. PAT Any Failure Rate, by 3-Month Follow-Up Period & Risk Level *(N = 729)* [a]

	n	Failure rate (%)
Risk Level		
Low	12	8.8
Moderate	64	23.4
High	153	47.8
Total	229	31.4
	$\chi^2 = 80.2^{***}$	

---

[27] Statistical significance was calculated with a chi-square test of association. Results are significant at the .01 alpha level (*p* < .001).
[28] Pearson's *r* correlation is statistically significant at the .01 alpha level (*p* < .001).

AUC-ROC, PAT total score	.72
Lower 95% CI	.68
Upper 95% CI	.75
Pearson's r	.34***
Logistic Regression	.07***

Note. PAT = Pretrial Assessment Tool; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval; Logistic regression results present marginal effects from a binomial logit model.
* *p* < .05, ** *p* < .01, *** *p* < .001
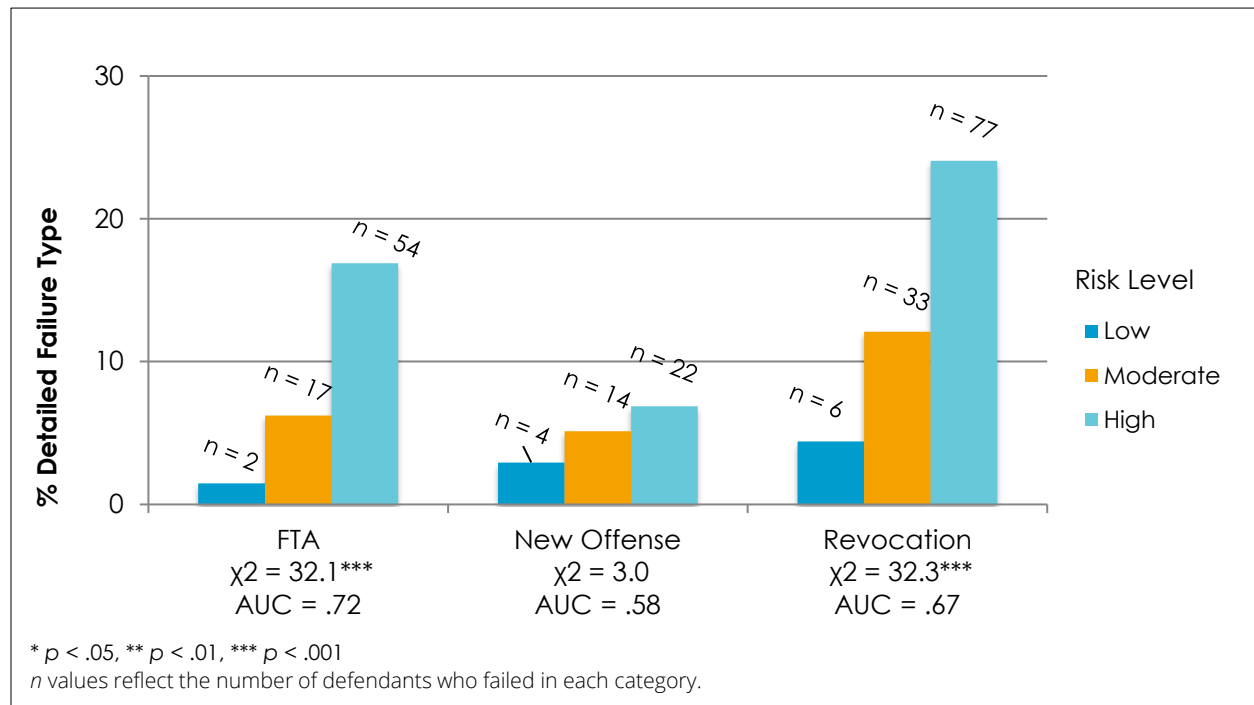[a.] 3-month follow-up sample totals: Low = 136; Moderate = 273; High = 320; Total = 729

Table 6 also includes predictive validity findings from logistic regression analysis, measuring the relationship between PAT total score and the likelihood of having a pretrial failure.[29] For ease of interpretation, logistic regression results are reported in terms of marginal effects (see **Appendix E** for complete logistic regression results). These results again indicate a strong statistically significant relationship between PAT total score and the likelihood of any pretrial failure. Specifically, the marginal effect results indicate that a one-point increase in PAT total score is associated, on average, with a seven-percentage point increase in the likelihood of any pretrial failure for released defendants.

Individual Failure Type Predictive Validity

While the PAT was designed to provide a comprehensive prediction estimate for a defendant's likelihood of either failing to appear and/or committing a new offense, this study went further by conducting additional predictive validity analyses. These analyses aimed to gauge the tool's individual predictive capabilities for FTAs, new offenses, and revocations. The following figure and table display the rates of each type of pretrial failure within three months by PAT risk level for the full sample of defendants (see **Figure 2** and **Table 7**). Defendants assessed in the low-risk category had a 1-4% failure rate for FTA, new offenses, or revocation in the three-month follow-up period. Defendants assessed in the moderate-risk category had a 5-12% pretrial failure rate within the same time frame for each failure type, while defendants assessed in the high-risk category had a 7-24% pretrial failure rate. As shown in **Figure 2,** the step-patterned increases in FTA, new offense, and revocation failure rates as risk levels increase from low to high align with our expectations.

---

[29] The validation sample was trimmed during data preparation to limit the pool of defendants to those who had been on pretrial release for at least 3 months and only captured pretrial failures occurring within 3 months (see **Methodology** for more detail). As a result, it was not necessary to control for time spent released in the logistic regression analysis. The outcome variables of pretrial failures and PAT total score were the only variables included in this analysis.

Figure 2. Detailed PAT Failure Rate, by 3-Month Follow-Up Period & Risk Level *(N = 729)*

* p < .05, ** p < .01, *** p < .001
n values reflect the number of defendants who failed in each category.

Additional testing displayed in **Table 7** found a statistically significant difference between risk levels and FTA and revocation failure rates, but not new offenses.[30] Confirming **Figure 2**'s visual trend of increasing pretrial failure rates moving from low- to high-risk, the calculated Pearson's *r* correlation coefficient showed fair predictive value that was statistically significant and positive for FTA failures (.23) and revocation failures (.21), meaning that the chance of each pretrial failure type increases as a defendant's assessed risk level increases.[31] The calculated Pearson's *r* correlation coefficient for new offenses was positive (.06), albeit weak and not statistically significant.

Following a similar pattern, validation analysis also found an excellent AUC-ROC value (.72) for FTAs, a good AUC-ROC value (.67) for revocation failures, and a fair AUC-ROC value (.58) for new offenses. These findings indicate that a defendant referred again to the Pretrial Services Unit for an FTA, new offense, or revocation three months after release would have a higher PAT score compared to a randomly selected defendant not referred again about 72%, 67%, and 58% of the time, respectively.

Logistic regression results measuring the relationship between PAT total score and the likelihood of each pretrial failure type are presented in **Table 7**.[32] For ease of interpretation, logistic regression results are reported in terms of marginal effects (see **Appendix E** for complete logistic regression results). These results again indicate a strong statistically significant relationship between PAT total score and the likelihood of an FTA or revocation. Specifically, the marginal effect results indicate

---

[30] Statistical significance was calculated with a chi-square test of association. Results are significant at the 0.01 alpha level (*p* < .001).

[31] Pearson's *r* correlation is statistically significant at the 0.01 alpha level (*p* < .001).

[32] The validation sample was trimmed during data preparation to limit the pool of defendants to those who had been on pretrial release for at least 3 months and only captured pretrial failures occurring within 3 months (see **Methodology** for more detail). As a result, it was not necessary to control for time spent released in the logistic regression analysis. The outcome variable of pretrial failures and PAT total score were the only variables included in this analysis.

that a one-point increase in PAT total score is associated, on average, with a three-percentage point increase in the likelihood of an FTA or revocation for released defendants.

Table 7. Percentage of Defendants with Failure within 3 months of Release, by Risk Level & Failure Type

	FTA [a]		New Offense [b]		Revocation [c]	
	n	Failure rate (%)	*n*	Failure rate (%)	*n*	Failure rate (%)
Risk Level						
Low	2	1.5	4	2.9	6	4.4
Moderate	17	6.2	14	5.1	33	12.1
High	54	16.9	22	6.9	77	24.1
Total	73	10.0	40	5.5	116	15.9
		$\chi^2 = 32.1$***		$\chi^2 = 3.0$		$\chi^2 = 32.3$***
AUC-ROC, PAT total score		.72		.58		.67
Lower 95% CI		.67		.49		.62
Upper 95% CI		.77		.66		.72
Pearson's r		.23***		.06		.21***
Logistic Regression		.03***		.01		.03***

Note. PAT = Pretrial Assessment Tool; FTA = failure to appear; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval; Logistic regression results present marginal effects from a binomial logit model.
* $p < .05$, ** $p < .01$, *** $p < .001$
[a.] FTA 3-month follow-up totals: Low = 37; Moderate = 104; High = 160; Total = 301
[b.] New Offense 3-month follow-up totals: Low = 61; Moderate = 95; High = 95; Total = 251
[c.] Revocation 3-month follow-up totals: Low = 23; Moderate = 49; High = 48; Total = 120

Race, Gender, Charge Level Predictive Validity

The preceding predictive validity analyses were calculated again across race, gender, and charge level to determine the extent to which the PAT correctly assesses risk levels for different sub-groups. Once again, and as expected, rates of any pretrial failure increased as risk level increased for every sub-analysis and produced good-excellent AUC-ROC values, indicating the PAT has predictive validity for defendants regardless of their race, gender, or highest booking charge level (see **Figures 3-5** and **Tables 8-9**).

These findings should be interpreted with caution because of small sub-analysis sample sizes by both risk-level and demographic variables. For the PAT race and charge level sub-analyses in particular, the sample size of defendants with an "Other" racial identity and low-risk level is just 15, while the sample size for the misdemeanor or infraction charge level and low-risk level is 16. As a result, observed pretrial failure rates are dependent on the outcomes of a limited number of defendants in each subgroup.[33]

Predictive Validity by Race

The following figure and table display the rates of any pretrial failure (i.e., FTA, new offense, or revocation) at three-month follow-up by PAT risk category and defendants' race (see **Figure 3**

---

[33] Note that only defendants' initial assessments were selected for inclusion in the analysis. As such, assessments for later referrals to the Pretrial Services Unit were removed to ensure each defendant was represented only once in each sample. Defendants without the requisite fixed follow-up duration were also removed.

and **Table 8**). Identical to previous figures presented, the step-patterned increase in pretrial failure displayed in **Figure 3** aligns with our expectations as the assessed risk level increases from low to high risk for each racial group. Specifically, the rate of any pretrial failure was as follows for White defendants: 14% (low), 23% (moderate), 44% (high). The pretrial failure rates were as follows for Hispanic defendants: 10% (low), 21% (moderate), 47% (high). Additionally, the pretrial failure rates were as follows for Black defendants: 4% (low), 27% (moderate), 60% (high). Lastly, the pretrial failure rates for defendants of an "Other" race were as follows: 0% (low), 28% (moderate), 47% (high).

Figure 3. PAT Any Failure Rates within 3 months of Release, by Risk Level & Race (*N* = 729)
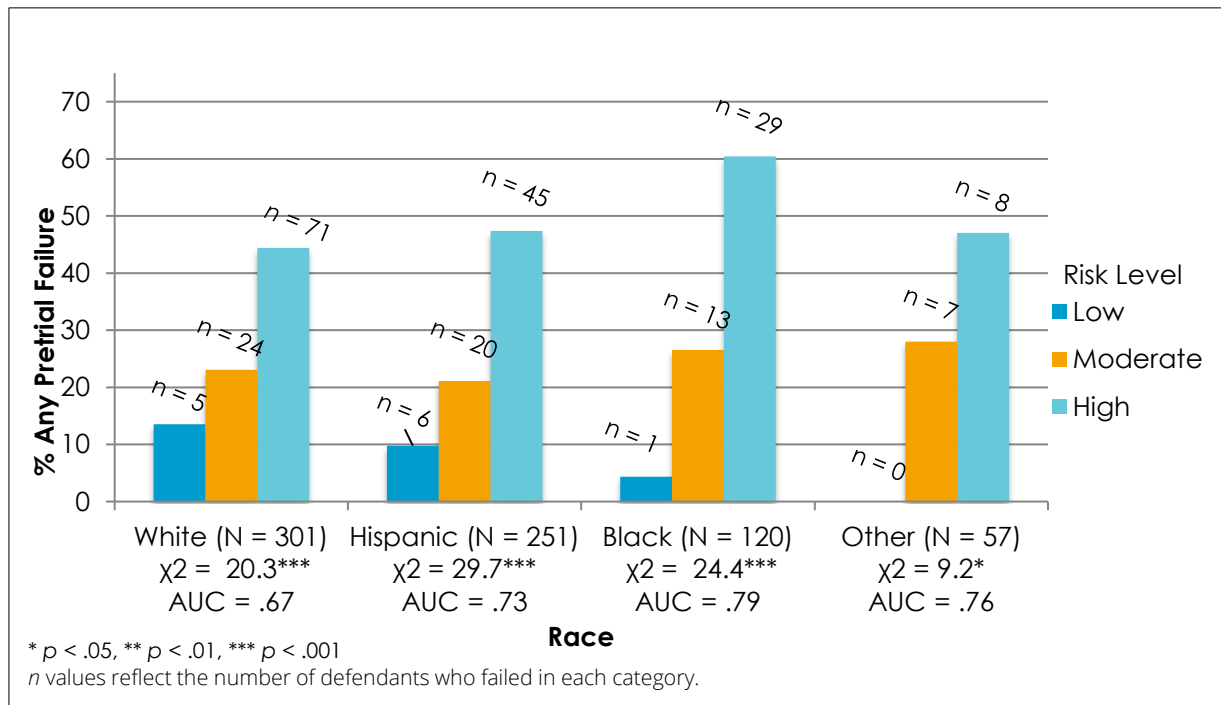
As expected, chi-square test found statistically significant differences between risk categories and any pretrial failure for each racial sub-group (see **Table 8**), meaning the observed differences in any pretrial failure rates are not due to chance alone for each racial identity.[34] In keeping with the clear visual trend in **Figure 3**, the calculated Pearson's *r* correlation coefficients showed good and excellent predictive values that were positive and statistically significant for each racial group analyzed as well.[35] This validation found excellent AUC-ROC values for assessed Hispanic defendants (.73), Black defendants (.79), and defendants of an "Other" race (.76). AUC-ROC values were good for White defendants (.67).

Table 8 also includes predictive validity findings from logistic regression analysis, measuring the relationship between PAT total score and the likelihood of having any pretrial failure for each

---

[34] Statistical significance was calculated with a chi-square test of association. Results are significant at the 0.01 alpha level for White, Hispanic, and Black defendants ($p < .001$). Results are significant at the 0.05 alpha level for defendants of an "Other" race ($p < .05$).

[35] Pearson's *r* is 0.27 (good) for White defendants, 0.36 (good-excellent) for Hispanic defendants, 0.48 (excellent) for Black defendants, and 0.40 (excellent) for defendants of an "Other" race. Pearson's *r* correlations are statistically significant at the 0.01 alpha level ($p < .01$) for each racial group.

racial group.[36] For ease of interpretation, logistic regression results are reported in terms of marginal effects. These results indicate there is a strong statistically significant relationship between PAT total score and the likelihood of any pretrial failure for each racial group. Specifically, the marginal effect results indicate that a one-point increase in PAT total score is associated on average with a seven-percentage point increase in the likelihood of any pretrial failure for White defendants, Hispanic defendants, and defendants of an "Other" race, respectively. Additionally, the marginal effect results indicate that a one-point increase in PAT total score is associated on average with an eight-percentage point increase in the likelihood of any pretrial failure for Black defendants.

Table 8. Percentage of Defendants with Any Pretrial Failure within 3 months of Release, by Risk Level & Race

	White [a]		Hispanic [b]		Black [c]		Other [d]	
	n	Failure rate (%)	*n*	Failure rate (%)	*n*	Failure rate (%)	*n*	Failure rate (%)
Risk Level								
Low	5	13.5	6	9.8	1	4.4	0	.0
Moderate	24	23.1	20	21.1	13	26.5	7	28.0
High	71	44.4	45	47.4	29	60.4	8	47.1
Total	100	33.2	71	28.3	43	35.8	15	26.3
		$\chi^2 = 20.3$***		$\chi^2 = 29.7$***		$\chi^2 = 24.4$***		$\chi^2 = 9.2$*
AUC-ROC, PAT total score		.67		.73		.79		.76
Lower 95% CI		.60		.67		.71		.63
Upper 95% CI		.73		.80		.87		.88
Pearson's r		.27***		.36***		.48***		.40**
Logistic Regression		.07***		.07***		.08***		.07***

Note. PAT = Pretrial Assessment Tool; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval; Logistic regression results present marginal effects from a binomial logit model.
* $p < .05$, ** $p < .01$, *** $p < .001$
[a.] White 3-month follow-up totals: Low = 37; Moderate = 104; High = 160; Total = 301
[b.] Hispanic 3-month follow-up totals: Low = 61; Moderate = 95; High = 95; Total = 251
[c.] Black 3-month follow-up totals: Low = 23; Moderate = 49; High = 48; Total = 120
[d.] "Other" 3-month follow-up totals: Low = 15; Moderate = 25; High = 17; Total = 57

The presented AUC confidence intervals in **Table 8** and complete logistic regression results displayed in **Appendix F (Table F1)** allow us to assess any difference in predictive validity between different racial groups. Each racial group has overlap in the 95% confidence intervals for the AUC (see **Table 8**), as a result, any observed differences in AUC scores may be due to chance. This finding that there are no racial differences in the predictive validity for any pretrial failure is supported by the logistic regression analysis. As shown in **Appendix F (Table F1),** being Hispanic, Black, or an "Other" race is not associated with statistically significant different pretrial failure rates compared to White defendants with the same PAT total score.

---

[36] The validation sample was trimmed during data preparation to limit the pool of defendants to those who had been on pretrial release for at least 3 months and only captured pretrial failures occurring within 3 months (see **Methodology** for more detail). As a result, it was not necessary to control for time spent released in the logistic regression analysis. The outcome variable of any pretrial failure, PAT total score, and a race variable were the only variables included in this analysis.

Predictive Validity by Gender & Highest Booking Charge Level

The following figure and table display the rates of any pretrial failure (i.e., FTA, new offense, or revocation) at three-month follow-up by PAT risk level, gender, and highest booking charge level (see **Figures 4-5** and **Table 9**). In line with previous findings and our expectations, **Figures 4-5** demonstrate a step-patterned increase in pretrial failure rates as male and female and misdemeanant and felon risk levels increase from low to high. Specifically, the rate of any pretrial failure was as follows for male defendants: 9% (low), 21% (moderate), 48% (high). The pretrial failure rates were as follows for female defendants: 7% (low), 32% (moderate), 47% (high). Additionally, the pretrial failure rates were as follows for defendants with a misdemeanor booking charge: 0% (low), 23% (moderate), 35% (high). Lastly, the pretrial failure rates for defendants with a felony booking charge were as follows: 11% (low), 24% (moderate), 51% (high).

Figure 4. PAT Any Failure Rate within 3 months of Release, by Risk Level & Gender *(N = 724)*
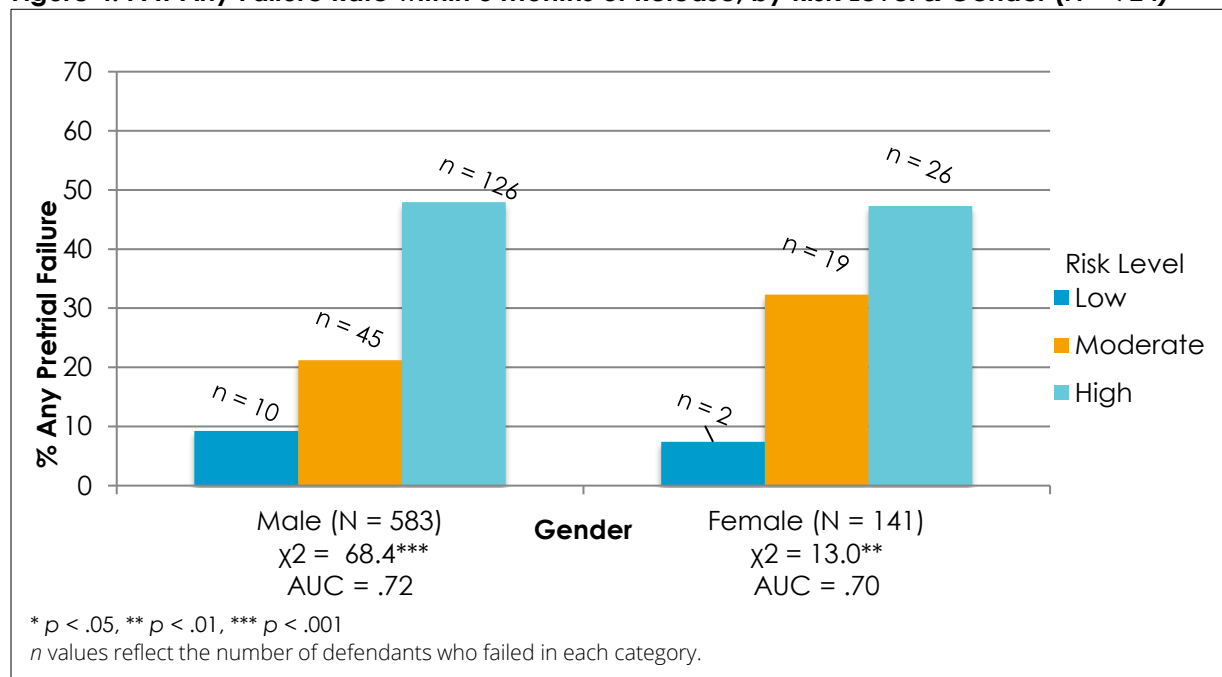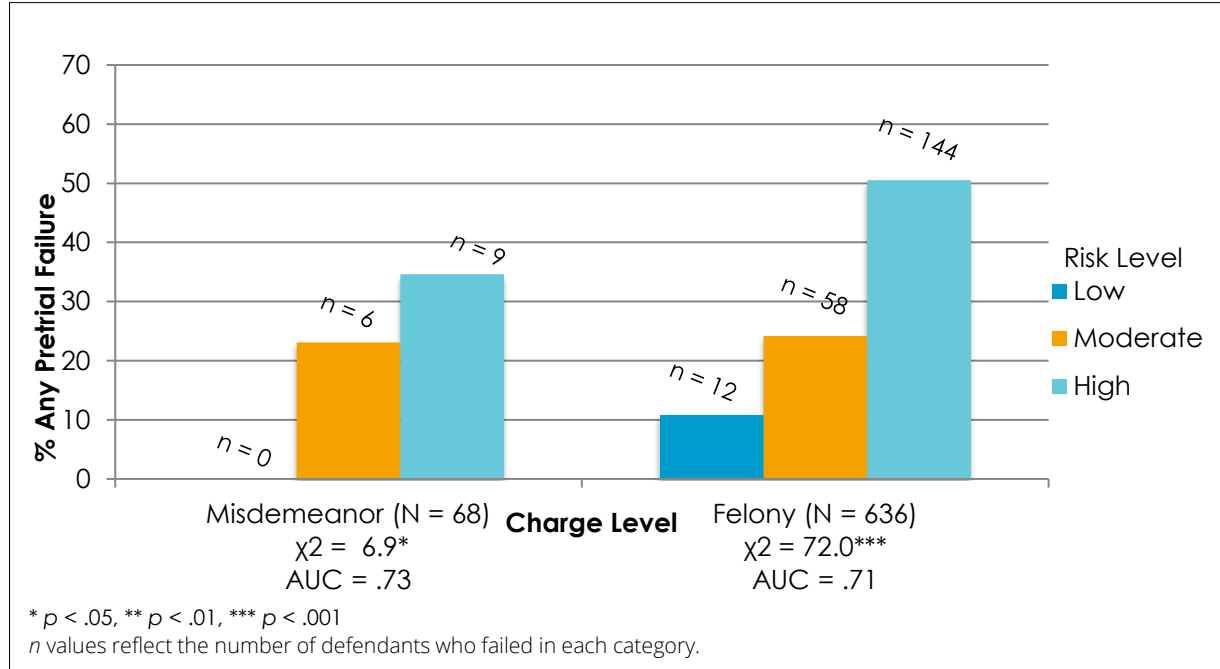

* *p* < .05, ** *p* < .01, *** *p* < .001
n values reflect the number of defendants who failed in each category.

Figure 5. PAT Any Failure Rate within 3 months of Release, by Risk Level & Highest Booking Charge Level (*N* = 704)

$* p < .05, ** p < .01, *** p < .001$
n values reflect the number of defendants who failed in each category.

As expected, chi-square tests found statistically significant differences between risk categories and any pretrial failure for gender and highest booking charge level (see **Table 9**), meaning the observed differences in any pretrial failure rates are not due to chance alone for each defendant gender identity or charge level.[37] In keeping with the clear visual trends in **Figures 4-5**, the calculated Pearson's *r* correlation coefficients showed a good predictive value that was positive and statistically significant for gender and charge level as well.[38] This validation found excellent AUC-ROC values for assessed male defendants (.72) and defendants with a misdemeanor or infraction booking charge (.73). AUC-ROC values were good-excellent for female defendants (.70) and defendants with a felony or enhancement booking charge (.71).

Logistic regression results measuring the relationship between PAT total score and the likelihood of any pretrial failure are presented in **Table 9**, broken out again by gender and highest booking charge level.[39] For ease of interpretation, logistic regression results are reported in terms of marginal effects (see **Appendix F** for complete logistic regression results). These results indicate again that there is a strong statistically significant relationship between PAT total score and any pretrial failure for each gender and charge level group. Specifically, the marginal effect results indicate that a one-point increase in PAT total score is associated on average with a seven-

---

[37] Statistical significance was calculated with a chi-square test of association. Results are significant at the .01 alpha level for male and female defendants ($p < .01$). Results are significant at the .05 alpha for defendants with misdemeanor/infraction or felony/enhancement charge levels ($p < .05$).

[38] Pearson's *r* is .35 (good) for male defendants, .33 (good) for female defendants, .33 (good) for defendants with a misdemeanor/infraction booking charge, and .34 (good) for defendants with a felony/enhancement booking charge. Pearson's *r* correlations are statistically significant at the .01 alpha level ($p < .01$) for each defendant group.

[39] The validation sample was trimmed during data preparation to limit the pool of defendants to those who had been on pretrial release for at least 3 months and only captured pretrial failures occurring within 3 months (see **Methodology** for more detail). As a result, it was not necessary to control for time spent released in the logistic regression analysis. The only variables included were an outcome variable for any pretrial failure, PAT total score, and either a gender variable or highest booking charge variable.

percentage point increase in the likelihood of any pretrial failure for male defendants, female defendants, and defendants with a felony booking charge, respectively. Additionally, the marginal effect results indicate that a one-point increase in PAT total score is associated on average with a six-percentage point increase in the likelihood of any pretrial failure for defendants with a misdemeanor booking charge.

Table 9. Percentage of Defendants with Any Pretrial Failure within 3 months of Release, by Risk Level, Gender, & Highest Booking Charge Level

| | Gender | | | | Highest Booking Charge Level | | | |
| | Male [a] | | Female [b] | | Misdemeanor or Infraction [c] | | Felony or Enhancement [d] | |
	n	Failure rate (%)	*n*	Failure rate (%)	*n*	Failure rate (%)	*n*	Failure rate (%)
Risk Level								
Low	10	9.3	2	7.4	0	.0	12	10.8
Moderate	45	21.2	19	32.2	6	23.1	58	24.2
High	126	47.9	26	47.3	9	34.6	144	50.5
Total	181	31.1	47	33.3	15	22.1	214	33.7
	$\chi^2 = 68.4^{***}$		$\chi^2 = 13.0^{**}$		$\chi^2 = 6.9^{*}$		$\chi^2 = 72.0^{***}$	
AUC-ROC, PAT total score	.72		.70		.73		.71	
Lower 95% CI	.68		.61		.61		.67	
Upper 95% CI	.76		.79		.85		.75	
Pearson's r	.35***		.33***		.33**		.34***	
Logistic Regression	.07***		.07***		.06***		.07***	

Note. PAT = Pretrial Assessment Tool; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval; Logistic regression results present marginal effects from a binomial logit model.
* $p < .05$, ** $p < .01$, *** $p < .001$
[a.] Male 3-month follow-up totals: Low = 108; Moderate = 212; High = 263; Total = 583
[b.] Female 3-month follow-up totals: Low = 27; Moderate = 59; High = 55; Total = 141
[c.] Misdemeanor 3-month follow-up totals: Low = 16; Moderate = 26; High = 26; Total = 68
[d.] Felony 3-month follow-up totals: Low = 111; Moderate = 240; High = 285; Total = 636

The presented AUC confidence intervals in **Table 9** and complete logistic regression results displayed in **Appendix F (Tables F2 & F3)** allow us to assess any difference in predictive validity between different gender identities and booking charge level. Each gender and booking charge level group has overlap in the 95% confidence intervals for the AUC (see **Table 9**), as a result, any observed differences in AUC scores may be due to chance. This finding that there are no gender or charge level differences in the predictive validity for any pretrial failure is supported by the logistic regression analysis. As shown in **Appendix F (Table F2),** being female as opposed to male is not associated with statistically significant different pretrial failure rates for defendants with the same PAT total score. Additionally, having a felony charge level as opposed to misdemeanor is not associated with statistically significant different pretrial failure rates for defendants with the same PAT total score (see **Appendix F, Table F3).**

III. Summary and Recommendations

This study used a mixed-methods approach to data collection and analysis to determine if the PAT is accurately predicting pretrial failures. The research team integrated staff perspectives with administrative data to paint a complete picture regarding assessment practices and implementation while also contextualizing predictive validity results. The study also set out to better understand and reduce any biases based on gender, race, or the highest booking charge level.

Local assessment practices promote PAT predictive accuracy. All officers have been trained to use the tool, administer the assessment in a face-to-face interview, refer to their interview guide, and collect collateral information. While officers did not refer to the scoring guide for each assessment, quality assurance practices already in place support accurate scoring. Crucially, officers are supportive of the tool, find that it is accurate, and are confident that scoring is consistent within the Pretrial Services Unit.

Ultimately, this study has found that the PAT has excellent predictive validity for any pretrial failure in Yolo County, as demonstrated by the high AUC-ROC values presented in **Table 10** and the statistically significant relationships identified between PAT total score and pretrial failure rate using chi-square, Pearson's *r*, and logistic regression.[40] **Furthermore, basic frequencies and statistical findings, such as AUC-ROC confidence intervals, indicate that the tool has excellent to good predictive validity and no detectable biases in the prediction of any pretrial failure by race, gender, or highest booking charge level.**

Table 10. Summarized PAT Predictive Validity Results, by Pretrial Failure Type

Failure Type	AUC-ROC, PAT total score	95% CI (Lower – Upper)
Any Failure*^	.72	.68 - .75
FTA*^	.72	.67 - .77
New Offense	.58	.49 - .66
Revocation*^	.67	.62 - .72

Note. PAT = Pretrial Assessment Tool; FTA = failure to appear; AUC-ROC = area under the receiver operating characteristic curve; CI = confidence interval; * Denotes statistically significant relationship between risk level (or total score) and failure using chi-square and Pearson's *r;* ^ Denotes statistically significant relationship between total score and failure using logistic regression.

The PAT's assessed predictive validity has become stronger over time in Yolo County, increasing for most measures of pretrial failure since the tool was last validated in 2021 (see **Table 11**). The predictive validity for FTAs improved most dramatically from a fair AUC-ROC value (.59) to an excellent value (.72). Similarly, predictive validity improved from good (.67) to excellent (.72) for any pretrial failure. The revocation AUC-ROC value improved modestly within a good predictive validity rating (.65 for 2016-2019, .67 for 2020-2023). The AUC-ROC predictive validity value declined for new offenses (.61 for 2016-2019, .58 for 2020-2023), although a .58 value is still considered fair predictive accuracy and does not change this study's overall finding that the PAT accurately predicts risk in Yolo County.

---

[40] Although the PAT was designed to predict any pretrial failure as opposed to FTA, new offenses, or revocation rates individually, predictive validity is also good for revocations and fair for new offenses.

Table 11. Comparison of PAT AUC-ROC Values Between Recent Validation Studies

Failure	AUC-ROC, PAT total score 2016-2019	AUC-ROC, PAT total score 2020-2023
Any Failure	.67	.72
FTA	.59	.72
New Offense	.61	.58
Revocation	.65	.67

Note. PAT = Pretrial Assessment Tool; AUC-ROC = area under the receiver operating characteristic curve; AUC-ROC for 2020-2023 applied exclusion criteria to sample that limited analysis to a period of three-months following release; AUC-ROC for 2016-2019 calculated by Bauman Consulting Group in 2021 and did not limit sample for analysis to a fixed period following release.

Recommendations

The improvement noted in the PAT's predictive validity since the last validation study is likely influenced by greater scoring consistency and accuracy as officers have gained more experience assessing defendants. To build on these strong predictive validity results and continue improving the reliability of the tool locally, the Yolo County Probation Department has already begun implementing some changes, such as creating a laminated scoring guide for officers to have available when scoring every PAT assessment. The County should also consider longer-term recommendations, such as developing a continuous quality improvement (CQI) team. A CQI team can support administration fidelity and implement strategies to improve inter-rater reliability and accuracy (e.g., consistent scoring between officers and agreement with tool master trainer scoring).

While PAT booster training is already planned for officers in 2024, a CQI team could regularly provide training opportunities to the Pretrial Services Unit between booster trainings and serve as coaches as PAT tool needs arise. These coaches (i.e., internal PAT trainers or supervisors) could observe officers while interviewing and scoring defendants using a structured feedback form that identifies areas of strength and areas for improvement and ongoing coaching as needed. These officer assessments could be done on a rotating basis to ensure each member of the Pretrial Services Unit receives feedback over the course of a year.

In particular, the Probation Department should consider a focus on scoring practices for drug-related questions (i.e., items six and seven) for its upcoming PAT booster training and CQI activities. Officers raised difficulty administering assessments with defendants who have severe drug abuse problems. Scoring for these items may also be driving higher average PAT total scores and weaker predictive accuracy for White defendants. While predictive accuracy was good for White defendants, it was lower than the excellent predictive accuracy Hispanic defendants, Black defendants, and defendants of an "Other" race.

As an additional component of CQI activities, continuous re-evaluation of the PAT at least every three years is recommended to comply with the requirements of SB 36 and to ensure the PAT continues to predict risk accurately. Yolo County should consider incorporating the same predictive validity analyses in future studies to measure specific changes in tool accuracy over time.

Acknowledgments

Appendix A: Yolo County ORAS-PAT Officer Survey

ORAS-PAT Administration & Use of Results[41]

1) How often do you administer the ORAS Pretrial Assessment Tool (ORAS-PAT)?*

( ) 5 or more times a week

( ) 1-4 times a week

( ) 2-3 times a month

( ) Once per month

( ) Less than once per month

( ) Other - Write In (Required): _____ *

2) Do you conduct interviews with the defendant during the ORAS-PAT assessment?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____ *

3) Do you have the ORAS-PAT interview guide in front of you during interviews?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____ *

4) Do you consult other sources of information when scoring the ORAS-PAT assessment (i.e., to verify the accuracy of defendant statements)? If yes, please provide examples.*

( ) Yes - Write In (Required): _____ *

( ) No

( ) Other - Write In (Required): _____ *

5) Do you reference the scoring guide when scoring an ORAS-PAT assessment?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____ *

---

[41] An asterisk (*) indicates a question is required.

6) Do you or your agency use the ORAS-PAT assessment information to assist in release decisions according to the release conditions framework?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____*

7) Does your agency use the assessment information to help determine pretrial supervision level?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____*

Quality Assurance & Release Decision Framework

8) Have you received formal training on how to use the ORAS-PAT?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____*

9) Have you received any retraining on the ORAS-PAT (i.e., attended a recertification or booster session)?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____*

10) If yes, how often do these trainings occur?*

( ) Once every six months

( ) Annually

( ) Once every two years

( ) Other - Write In (Required): _____*

11) Is there a quality assurance process (e.g., supervisor approval or having someone observe your interview and scoring) in place concerning the ORAS-PAT?*

( ) Yes

( ) No

( ) Other - Write In (Required): _____*

12) In addition to the ORAS-PAT risk level, what other pieces of information do you consider as you're making your release recommendation?*

_____

13) How do you prioritize these pieces of information (e.g., ORAS-PAT, instant charge, victim statement) when developing a release recommendation (i.e., what carries more weight)?*

_____

ORAS-PAT Fairness, Benefits, & Staff Support

14) Please indicate your level of agreement with the following statements related to the ORAS-PAT as it is used by your agency.*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Uncertain
The ORAS-PAT has helped our agency establish a common language regarding defendants' risk level	( )	( )	( )	( )	( )	( )
The ORAS-PAT is easy to read, interpret, and use	( )	( )	( )	( )	( )	( )
The ORAS-PAT increases consistency in rules, policies, and procedures regarding assessment practices	( )	( )	( )	( )	( )	( )
The ORAS-PAT reduces variation in decision-making across individual staff	( )	( )	( )	( )	( )	( )

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Uncertain
The ORAS-PAT helps our agency make proper decisions regarding whether a defendant is appropriate for pretrial supervision versus custody	( )	( )	( )	( )	( )	( )
The ORAS-PAT reduces bias in criminal justice decision making in our agency	( )	( )	( )	( )	( )	( )

15) Please indicate your level of agreement with the following statements related to *usefulness* the ORAS-PAT at your agency.*

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Uncertain
The ORAS-PAT assessment has made my job easier	( )	( )	( )	( )	( )	( )
Risk and needs assessment benefits defendants	( )	( )	( )	( )	( )	( )
There is staff support for implementing the ORAS-PAT	( )	( )	( )	( )	( )	( )
There are clear guidelines for when to use the ORAS-PAT	( )	( )	( )	( )	( )	( )
There is a protocol for how to use the ORAS-PAT	( )	( )	( )	( )	( )	( )

16) How confident are you that the score produced by the ORAS-PAT accurately captures the likelihood of a defendant recidivating.*

( ) Completely lack confidence

( ) Not confident

( ) Neutral

( ) Confident

( ) Completely confident

( ) Uncertain - Write In (Required): _____ *

17) How confident are you that a defendant will receive the same score no matter which staff member conducted the assessment.*

( ) Completely lack confidence

( ) Not confident

( ) Neutral

( ) Confident

( ) Completely confident

( ) Uncertain - Write In (Required): _____ *

18) How do you rate your overall satisfaction with the ORAS-PAT?*

( ) Very Dissatisfied

( ) Dissatisfied

( ) Neutral

( ) Satisfied

( ) Very Satisfied

( ) Other - Write In (Required): _____ *

ORAS-PAT Strengths & Challenges

19) List 3 challenges of using the ORAS-PAT assessment in your agency. This question is required, please write "NA" when you cannot think of a challenge.*

Challenge 1: _____

Challenge 2: _____

Challenge 3: _____

20) List 3 strengths of using the ORAS-PAT assessment in your agency. This question is required, please write "NA" when you cannot think of a strength.*

Strength 1: _____

Strength 2: _____

Strength 3: _____

Closing Thoughts

21) Is there anything else you would like to share with us?

_____

_____

Staff Information

22) What is your highest level of education completed?*

( ) GED

( ) High school diploma

( ) Associate's degree

( ) Bachelor's degree

( ) Graduate degree

23) How many years of experience do you have working in the criminal justice field (including your time at probation)?*

_____

24) How long have you been working at probation?*

_____

25) How long have you been working in your current role in the Pretrial Unit?*

_____

26) How many defendants are currently on your caseload?*

_____

Thank You!

Appendix B: Detailed Defendant Characteristics, by Pretrial Release Decision

Table B1. Assessed Defendant Booking Charge Characteristics, by Risk Level

	Overall (n = 704)		Misdemeanor or Infraction (n = 68)		Felony or Enhancement (n = 636)	
	n	%	*n*	%	*n*	%
Risk Level						
Low	217	16.3%	16	23.5%	111	17.5%
Moderate	481	36.2	26	38.2	240	37.7
High	630	47.4	26	38.2	285	44.8

* p < .05, ** p < .01, *** p < .001

Table B2. Assessed Defendant Characteristics, by Release Decision (N = 1,622)

	Overall (N = 1,622)		Denied & Other (N = 893, 55%)		Released (N = 729, 45%)	
	n	%	*n*	%	*n*	%
Gender						
Male	1,303	81.2%	720	81.8%	583	80.5%
Female	301	18.8	160	18.2	141	19.5
Race						
White	641	39.5%	340	38.1%	301	41.3%
Hispanic	533	32.9	282	31.6	251	34.4
Black	310	19.1	190	21.3	120	16.5
Other	138	8.5	81	9.1	57	7.8
Booking Charge Level**						
Misdemeanor or Infraction	165	12.4%	97	15.5%	68	9.7%
Felony or Enhancement	1,163	87.6	527	84.5	636	90.3
Risk Level						
Low	306	18.9%	170	19.0%	136	18.7%
Moderate	591	36.4	318	35.6	273	37.4
High	725	44.7	405	45.4	320	43.9

* p < .05, ** p < .01, *** p < .001

Appendix C: Detailed Defendant Characteristics, by Probation Release Recommendation Agreement with Judicial Officer

Table C1. Assessed Defendant Characteristics, by Probation Release Recommendation Agreement with Judicial Officer (N = 319)

	Overall (*N* = 319)		Agreement (*N* = 201, 63%)		Disagreement (*N* = 118, 37%)	
	n	%	*n*	%	*n*	%
Gender**						
Male	242	78.1%	161	83.0%	81	69.8%
Female	68	21.9	33	17	35	30.2
Race						
White	121	37.9%	76	37.8%	45	38.1%
Hispanic	110	34.5	62	30.8	48	40.7
Black	50	15.7	32	15.9	18	15.3
Other	38	11.9	31	15.4	7	5.9
Booking Charge Level						
Misdemeanor or Infraction	27	10.8%	14	9.5%	13	12.7%
Felony or Enhancement	222	89.2	133	90.5	89	87.3
Risk Level						
Low	78	24.5%	51	25.4%	27	22.9%
Moderate	119	37.3	74	36.8	45	38.1
High	122	38.2	76	37.8	46	39
Release Decision***						
Denied	158	49.5%	144	71.6%	14	11.9%
Released	161	50.5	57	28.4	104	88.1

* *p* < .05, ** *p* < .01, *** *p* < .001

Appendix D: Detailed Defendant Characteristics, by Pretrial Success or Failure

Table D1. Released Defendant Characteristics, by Pretrial Success or Failure within 3 months

	Overall (n = 729)		Successful (n = 500, 69%)		Any Failure (n = 229, 31%)	
	n	*%*	*n*	*%*	*n*	*%*
Gender						
Male	583	80.5%	402	81.0%	181	79.4%
Female	141	19.5	94	19	47	20.6
Race						
White	301	41.3%	201	40.2%	100	43.7%
Hispanic	251	34.4	180	36	71	31
Black	120	16.5	77	15.4	43	18.8
Other	57	7.8	42	8.4	15	6.6
Booking Charge Level						
Misdemeanor or Infraction	68	9.7%	53	11.2%	15	6.6%
Felony or Enhancement	636	90.3	422	88.8	214	93.4
Risk Level*						
Low	136	18.7%	124	24.8%	12	5.2%
Moderate	273	37.4	209	41.8	64	27.9
High	320	43.9	167	33.4	153	66.8

* p < .05, ** p < .01, *** p < .001

Appendix E: Logistic Regression Results, by Pretrial Failure Type

Table E1. Binomial Logit Model PAT Total Score & Any Failure, 3-Month Follow-Up Period

	Any Failure	Any Failure (ME)
Total Score	.37***	.07***
	(.04)	(.01)
Constant	-2.68***	
	(.25)	
Observations	729	729

Note. ME = marginal effects.
Standard errors in parentheses
* p < .05, ** p < .01, *** p < .001

Table E2. Binomial Logit Model PAT Total Score & Detailed Failure Type, 3-Month Follow-Up Period

	FTA	FTA (ME)	New Offense	New Offense (ME)	Revocation	Revocation (ME)
Total Score	.39***	.03***	.12	.01	.27***	.03***
	(.07)	(.01)	(.07)	(.00)	(.05)	(.01)
Constant	-4.39***		-3.44***		-3.12***	
	(.44)		(.43)		(.30)	
Observations	729	729	729	729	729	729

Note. ME = marginal effects; FTA = failure to appear.
Standard errors in parentheses
* p < .05, ** p < .01, *** p < .001

Appendix F: Logistic Regression Results for Any Pretrial Failure, by Defendant Characteristics

Table F1. Binomial Logit Model PAT Total Score & Any Pretrial Failure in 3-Month Follow-Up Period, Race

	Any Fail	Any Fail (ME)
Total Score	.37***	
	(.04)	
Race: Hispanic	.02	
	(.20)	
Race: Black	.41	
	(.24)	
Race: "Other"	.16	
	(.35)	
Race: White (ME)		.07***
		(.01)
Race: Hispanic (ME)		.07***
		(.01)
Race: Black (ME)		.08***
		(.01)
Race: "Other" (ME)		.07***
		(.01)
Constant	-2.81***	
	(.28)	
Observations	729	729

Note. ME = marginal effects.
Standard errors in parentheses
$^*p < .05$, $^{**}p < .01$, $^{***}p < .001$

Table F2. Binomial Logit Model PAT Total Score & Any Pretrial Failure in 3-Month Follow-Up Period, Gender

	Any Failure	Any Failure (ME)
Total Score	.37*** (.04)	
Gender: Female	.26 (.21)	
Gender: Male (ME)		.07*** (.01)
Gender: Female (ME)		.07*** (.01)
Constant	-2.74*** (.26)	
Observations	724	724

Note. ME = marginal effects.
Standard errors in parentheses
* *p* < .05, ** *p* < .01, *** *p* < .001

Table F3. Binomial Logit Model PAT Total Score & Any Pretrial Failure in 3-Month Follow-Up Period, Charge Level

	Any Failure	Any Failure (ME)
Total Score	.37*** (.04)	
Charge: Felony or Enhancement	.54 (.32)	
Charge: Misdemeanor or Infraction (ME)		.06*** (.01)
Charge: Felony or Enhancement (ME)		.07*** (.01)
Constant	-3.13*** (.39)	
Observations	704	704

Note. ME = marginal effects.
Standard errors in parentheses
* *p* < .05, ** *p* < .01, *** *p* < .001